

NATIONAL SCIENCE FOUNDATION
4201 WILSON BOULEVARD
ARLINGTON, VIRGINIA 22230



OFFICE OF
INSPECTOR GENERAL

OIG Report: M90070030

I.

Introduction

This case involves [] (the "subject"), an [] in the [] at [] (the "institution"). The subject's proposal [] (the "proposal"), a renewal of NSF award [], was submitted to the [].

We received two allegations regarding differences between certain data reported in the proposal and results reported by the subject in an earlier article, [

[], authored by the subject, [] (the "postdoc"), [], [], [] (the "second graduate student"), [] (the "first graduate student"), [] (the "undergraduate"), and [] (the "article"). The article acknowledges financial support from NSF award [].

II.

The Institution's Actions

A.

Deferral

On 25 July 1991, we called the institution's designated Authorized Organizational Representative and informed him that we had received allegations of misconduct in science regarding one of the institution's faculty. We asked if the institution would prefer for this office to defer to them to conduct an inquiry, and investigation if necessary; we also told him that, because the allegations involved possible fabrication of research data, the institution should move expeditiously to secure the primary data in the subject's laboratory. The institution agreed to secure promptly the primary data and initiate an appropriate inquiry.

B.
The Initial "Inquiry"¹

We asked the institution to address two allegations; see Appendix I. We received a letter from the institution, dated 17 October 1991 (the "letter"), along with a "Report of the Committee of Inquiry" dated 18 September 1991 (the "report"), a written response by the subject, and copies of laboratory notebook pages and spectra related to the report. The letter said:

"The appropriate university committee has now concluded their inquiry on the questions raised in your letter. . . . The inquiry committee reports that there exists no evidence of academic misconduct with respect to the allegations stated in your letter. We do find however that [the subject] does need to exert more diligence in the manner he references his data, particularly when discrepancies or errors are being corrected."

The committee found that there were laboratory records that supported the experimental results that had been reported in the article. The subject was able to explain why certain experiments had been repeated, revealing that the results published in the article were in error:

"In summary, the data are all here that account for the published and submitted reports, and the corrections in the literature were undertaken prior to receipt of the NSF allegations by [the institution]. The problem of the NSF proposal was a sin of omission on [the subject]'s part by not stating that new data were employed in the Preliminary Results section; nothing sinister was intended. We conclude that there is no substance to the allegations brought against [the subject]."

¹Under NSF's misconduct regulation,

"An 'inquiry' consists of preliminary information-gathering and preliminary fact-finding to determine whether an allegation or apparent instance of misconduct has substance. An investigation must be undertaken if the inquiry determines the allegation or apparent instance of misconduct has substance. An 'investigation' is a formal examination and evaluation of relevant facts to determine whether misconduct has taken place or, if misconduct has already been confirmed, to assess its extent and consequences or determine appropriate action."

45 C.F.R. § 689.1(c). It is clear that what the institution conducted here was not mere "preliminary information-gathering and preliminary fact-finding" and in fact amounted to an investigation.

C.
Unresolved Issues

NSF's definition of "misconduct" includes "fabrication, falsification, plagiarism, or other serious deviation from accepted practices in proposing, carrying out, or reporting results from activities funded by NSF" ² It was clear from the report and its accompanying documentation that this was not a case of data fabrication or falsification. However, two NSF experts in the subject's scientific discipline reviewed the documentation, and they questioned whether some of the quantitative data the subject possessed was of sufficient quality to justify its reporting in either the article or the proposal.

Whether a particular scientist's practices constitute "serious deviation from accepted practices," under NSF's definition of misconduct in science, is a question that must properly be addressed by that scientist's community. We therefore wrote to the institution on 29 January 1992 and asked the institution to address eight specific issues concerning the adequacy of the subject's data; see Appendix II. Specifically, we asked the institution to advise us whether any of these issues, in the institution's opinion, constituted "serious deviation from accepted practices in proposing, carrying out, or reporting results from" research in the relevant scientific discipline.

D.
Further "Inquiry"

We received a letter from the institution, dated 19 February 1992 (the "second letter"), along with a "Second Report of the Committee of Inquiry" dated 18 February 1992, and a second written response by the subject (the "subject's second response"). In the second letter, the dean said:

"In reviewing the matter to greater depth in the areas raised in your most recent letter, the [institution's] inquiry committee continues to find that, in our judgment, [the subject] had no intent to deceive, falsify, fabricate, or plagiarize results, and that there exist verifications that experiments to substantiate the results quoted were indeed carried out by [the subject] and the various members of his group. Thus, we do not detect anything that is unethical or illegal.

"Now, we do report that there exist some less than ideal experimental observations in that work, and those do raise questions with respect to the validity of the data cited in the [article]. . . . It is clear here that, if there had been a careful monitoring of the experimental results, there would have resulted remeasurements of what are now very clearly questionable results reported in the [article].

²45 CFR § 689.1(a)(1).

"Now, the committee does report that there are some less than ideal procedures in [the subject]'s laboratory environment and in his style of leadership that do require corrective actions, but we consider these to be the responsibility of the [institution]. As such, they are being addressed by [the chairperson of the subject's department]. In particular, the committee has insisted that a set of 'good laboratory practices' be developed that [the subject] will be expected to review with all students and laboratory personnel henceforth.

"In closing, the [institution's] inquiry committee finds that there exist some irregularities associated with the experimental results being questioned, but these do not constitute misconduct as defined by the National Science Foundation.

..."

In the subject's second response, he acknowledged

"that certain laboratory experiments involving the work in question were performed poorly and that experiments should have been repeated when questions or deficiencies were either present or suspected. However, I categorically deny that I knowingly engaged in any scientific, ethical, or legal misconduct.

"What I failed to do was to maintain high standards of scientific work and reporting. . . . I regret these mistakes and state emphatically that mistakes such as these will not occur again.

"In the future I shall take appropriate precautions in reviewing and discussing all the data shown to me by my students and coworkers before I use their data in preparing manuscripts for publication and research proposals. I shall insist that all research observations are well-documented, that all calculations have been verified, that no conclusions are based on any assumptions, that all compounds are pure before any characteristics are assigned to them, and that all results are properly recorded in laboratory notebooks. Furthermore, I will insist that all coauthors carefully examine the data that appear in our drafts of publications and research proposals, and verify that the data in the final manuscripts and write-ups are the same as those reported in their notebooks. However, I realize the ultimate responsibility for the correctness of all data, results, and conclusions that I transmit is mine."

III. Conclusion

In light of the institution's responses in this matter, we conclude that there is not sufficient evidence to establish that the subject engaged in misconduct in science under NSF's regulation. Although we were concerned about some of the research practices documented in this matter, we accept the institution's judgment that those practices do not constitute "serious

deviation from accepted practices in proposing, carrying out, or reporting results from activities funded by NSF," under NSF's definition of misconduct in science. We are also reassured by the subject's response and the preventative actions taken by the subject and the institution, that the poor research practices that led to these allegations will not be repeated.

[REDACTED]

31 March 1992

Appendix I

Allegations

We asked the institution to address two allegations, as follows³:

"1. The proposal states on page 6 that [

].

"These results appear to be inconsistent. What data supported the report in the article of [], and what data support the different report in the proposal?"

"2. The article reports that [

].

"These results appear to be inconsistent. What data supported the report in the article of [] and what data support the different report in the proposal?"

³Citations to the scientific literature are omitted. The numerical designations for chemical structures in the proposal are referred to with a "p" preceding the number, and those in the article are preceded by an "a".

Appendix III

Additional Issues

We asked the institution to address eight specific issues concerning the adequacy of the subject's data, as follows⁴:

- "1. On page 4 of his letter, [the subject] explained the article's report of [] as follows:

"Starting from 0.5 g of [], 0.52 g of crude product was obtained. From column chromatography, [the first graduate student] isolated 0.32 g of material, but he had told me, and I have also written down in his notebook (in the presence of [the first graduate student and the postdoc]), that he had accidentally thrown away some fractions from column chromatographic separation which was estimated to be about 0.2 g (see [XX]-II-108) (since the original weight was 0.52 g, and he had concentrated all other fractions from column chromatography except those fractions that he had thrown away and found nothing).'

"It is our understanding that it is typical for some material to be left behind on a chromatographic column — particularly when the TLC shows a substantial spot remaining at the origin even with a fairly polar solvent, such as [the first graduate student]'s did.⁵ Therefore, we do not understand the basis for [the subject]'s presumption that the unknown quantity of eluent that was 'lost' contained the full 0.2 g difference between what was put onto the column and what came off.

- "2. [The subject] continued:

"I was told by [the first graduate student] that the yield was about 85% and this figure was reported. Probably, this number is derived from the analysis of the NMR spectrum (NMR [XX]-2-106-1) of the mixture (0.32 g; [XX]-2-107). The spectrum indicated a 5:1 ratio of the desired product [] (these two were separated and identified individually later).'

⁴The individual researcher's initials, which precede all notebook, compound, and spectra references, are replaced by "XX"; the numerical designations for chemical structures in the proposal are referred to with a "p" preceding the number, and those in the article are preceded by an "a".

⁵"See notebook page [XX]-2-107, middle of the page — '100% EA' presumably indicates the use of pure ethyl acetate: the spots at the far left presumably are a sample of the crude product that was placed on the column."

"It is our understanding that a reported yield generally indicates the amount of material isolated with a reasonable degree of purity. It appears from [the first graduate student]'s notebook that only 0.15 g of [] was actually isolated, which is a yield of 30%.⁶

"3. In his discussion of [the postdoc]'s work on page 5 of his letter, [the subject] states that the []

"'is [] indicating that [the first graduate student]'s early sample of [] was impure.'

"According to [the first graduate student]'s notebook, [] was measured on the material that had been purified, the NMR spectrum of which seems to indicate that it was at least as pure as [the postdoc]'s.⁷

"4. After discussing [the postdoc]'s work, on pages 5-6 of his letter [the subject] continued:

"'Looking back at [the first graduate student]'s data, it is apparent that 0.32 g material is correct. Most likely, the other 0.2 g of [] (which as explained by [the first graduate student] was lost by accident) was insoluble material in CH₂Cl₂ and stayed undissolved on the top of the chromatography column'

"As we understand it, the material 'lost by accident' had been eluted from the column, and therefore must have been soluble in whatever solvent ratio carried the material off the column (which is not specified in the notebook).⁸ Was the material lost because it was insoluble and stayed at the top of the column, or because it came off the column and was thrown away? In light of [the postdoc]'s findings regarding the low solubility of [], how justified was [the subject]'s 'estimate' that the eluent that had been thrown away contained 0.2 g of the desired product, as related in (a) above?

"5. The article stated that 'the syn disposition at [] is supported by ¹H NMR 2D NOESY and 2D COSY experiments.' Neither your report nor [the subject]'s letter

⁶"See notebook page [XX]-2-108 and NMR [XX]-2-106-3."

⁷"See notebook page [XX]-2-108; compare NMR [XX]-2-106-3 with NMR [XX]-5-91-2."

⁸"See notebook page [XX]-2-108: 'lost: 0.20 g (accident, throw away some fraction containing the compound)', and, as quoted above from page 4 of [the subject]'s letter, [the first graduate student] 'had accidentally thrown away some fractions from column chromatographic separation'."

addresses this issue.⁹

"6. [The subject] stated on page 6 of his letter that

"'As discussed in reply to allegation (1), compound [] was obtained by using []. . . . Since the product, [], was indeed (later found to be) a mixture of two compounds ([] and [], but contained greater amounts of []), the final product ([]) had a [] which was stated in the notebook ([XX]-I-53).'

"Although it is difficult to discern from the notebooks because none of the students or postdocs recorded the source of their starting materials, it appears that purified starting materials were used and purified products were obtained in each step.¹⁰ Therefore, the explanation quoted above for the low rotation obtained for the product the first time — before [the subject] was aware of the error in the reported rotation for [] — seems untenable.

"7. On page 8 of his letter, [the subject] explained the yield calculation for [] as follows:

"'[] (30 mg) was [] and 30 mg . . . of [] was obtained. NMR spectrum of this sample ([XX]-I-53-2) indicated a contamination of ~13% of the 30 mg material (about 4 mg) of solvents (ethyl acetate, MeOH and small impurity). This was then calculated to about 26 mg (93%) of [].'

"This calculation does not appear in the notebook. Please address the propriety of this method of calculating a yield, particularly a yield to be reported in a journal publication, and particularly when the 'contamination' is nothing more than volatile solvents that could be readily removed in vacuo.

⁹"There is what appears to be a 2D COSY spectrum among the materials you provided, but it is dated 26 May 1989 — before [the first graduate student] performed the reaction on [] June 1989 to make []. There is no NOESY spectrum."

¹⁰"Presumably, [the first graduate student] used as starting material for [] reaction 0.12 g of the 0.15 g of [] that he had just purified (see [XX]-2-108, NMR [XX]-2-106-3, and [XX]-2-111). The product from this reaction, [], was apparently purified by [the undergraduate], who isolated from column chromatography fractions 46-48 27.1 mg of [] shown by ¹H and ¹³C NMR to be quite pure (see [XX]-1-61, ¹H NMR [XX]-1-61-1, and ¹³C NMR [XX]-1-62-1). That was apparently added to the 34.1 mg of material from fractions 49-55, which according to the TLC should have been even more pure than fractions 46-48), to give 61.2 mg of []. [the second graduate student] apparently then carried out the [] on 30 mg of this [], and he purified the product [] by PTLC (see [XX]-1-53)."

"8. On page 10 of his letter, [the subject] explained the [] and yield calculations for [] as follows:

"'[XX]-102 (0.1 g) was [] to give 73 mg (88% yield) of []. [] was measured on Fraction 13-15 (56 mg with NMR [XX]-5-102-6) in which its NMR spectrum showed slight impurities and ethyl acetate solvent, responsible for a lower value than the reported value of [] [The postdoc] told me that if an allowance was given for the impurity, the [] would be around [] and the yield would be ~ 80% (instead of 88%).'

"On page 7 of his letter, [the subject] noted that [the postdoc] had [] a sample 'containing slight impurity and solvent', even though a 'slightly purer' sample was available which 'was not used in measuring []'.

"The calculations employed to arrive at the figures reported do not appear in the notebook. Please address the propriety of this method of estimating [] and yield, particularly when they are to be reported in a journal publication or grant proposal, particularly when much of the contamination is easily-removable volatile solvent, and particularly when a purer sample is available.^{11"}

¹¹"One page 2 of its report, your inquiry committee went through an elaborate process to estimate that the amount of ethyl acetate shown in NMR [XX]-5-106-2 could, '[a]ssuming a 50% error' in the measurements, account for 'over one-half of the 4.4 mg contamination cited by [the subject and the postdoc] present in the original 20 mg sample of [] used in the [] determination.' We do not see how this estimation allows you to evaluate whether [the subject] was justified with the data available to him in reporting to NSF that he had obtained an 80% yield of []."